

A Simple and Rapid System for Proteomic Analysis of the Archaeon *Candidatus Vulcanisaeta moutnovskia*

Nikolay A. Chernyh

Winogradsky Institute of Microbiology, Federal Research Center of Biotechnology, Russian Academy of Sciences

Sinje Neukirchen

Division of Archaea Biology and Ecogenomics, Department of Ecogenomics and Systems Biology, University of Vienna

Evgenii N. Frolov

Winogradsky Institute of Microbiology, Federal Research Center of Biotechnology, Russian Academy of Sciences

Filipa L. Sousa

Division of Archaea Biology and Ecogenomics, Department of Ecogenomics and Systems Biology, University of Vienna

Margarita L. Miroshnichenko

Winogradsky Institute of Microbiology, Federal Research Center of Biotechnology, Russian Academy of Sciences

Alexander Y. Merkel

Winogradsky Institute of Microbiology, Federal Research Center of Biotechnology, Russian Academy of Sciences

Nikolay V. Pimenov

Winogradsky Institute of Microbiology, Federal Research Center of Biotechnology, Russian Academy of Sciences

Dimitry Y. Sorokin

Winogradsky Institute of Microbiology, Federal Research Center of Biotechnology, Russian Academy of Sciences

Sergio Ciordia

Proteomics Facility, Centro Nacional de Biotecnología, CSIC

María C. Mena

Proteomics Facility, Centro Nacional de Biotecnología, CSIC

Manuel Ferrer (✉ mferrer@icp.csic.es)

Institute of Catalysis, CSIC <https://orcid.org/0000-0003-4962-4714>

Peter N. Golyshin

School of Biological Sciences, Bangor University

Alexander V. Lebedinsky

Winogradsky Institute of Microbiology, Federal Research Center of Biotechnology, Russian Academy of Sciences

Inês A. Cardoso Pereira

Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa

Elizaveta A. Bonch-Osmolovskaya

Winogradsky Institute of Microbiology, Federal Research Center of Biotechnology, Russian Academy of Sciences; Lomonosov Moscow State University, Faculty of Biology

Method Article

Keywords: protein extraction, archaea, *Candidatus Vulcanisaeta moutnovskia*

Posted Date: July 22nd, 2020

DOI: <https://doi.org/10.21203/rs.3.pex-1032/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

This protocol describes a rapid protein extraction method for the archaeon *Candidatus* *Vulcanisaeta moutnovskia*, which can be also implemented for other archaea. The utilization of two different methods for protein extraction constitute the main step of the protocol. Method I involves the extraction with a multi-chaotropic lysis buffer containing a non-denaturing zwitterionic detergent, most efficient for extracting cytosolic proteins. Method II involves a denaturing anionic detergent allowing total disruption of the membranes and capable of extracting both membrane (hydrophobic) and non-membrane (water-soluble, hydrophilic) proteins. The big advantage of the methods is to use general laboratory chemicals to make powerful extraction buffers, resulting in high quality and quantity of proteins. The methods probably are usable for any other archaea or microbial cells, and takes about 14-22 h. Following extraction and further protein digestion, 1D-nano Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometric (LC ESI-MSMS) analysis with Triple TOF 5600 and Orbitrap technologies were used for protein identification and further quantification.

Introduction

Gene expression is routinely quantified by measuring mRNA levels. However, there is uncertainty about how closely levels of mRNAs relate to levels of their corresponding proteins; several papers suggest only a modest correlation using transcriptomic and proteomic techniques^{1,2}. This is why the analysis of proteins that are actively synthesized by a microbe is now extensively recognized, as last step of the functional hierarchy, to provide associations with functions that become part of the active metabolism. Many protocols have been devised for extracting and analyzing proteins from diverse samples. However, case by case adaptations of those protocols are needed, in order to guarantee a correct representation of cytosolic and membrane bound proteins, the last one having common and inherent problems of extractions. Here we developed a two-step extraction method that can be universally used. It is based on the utilization of a first “soft” lysis buffer containing 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate (CHAPS) to extract mostly cytosolic proteins. membrane-bound proteins. Second, to improve the extraction efficiency of membrane proteins, we additionally used a new extraction method in which a multi-chaotropic lysis buffer containing 7 M urea, 2 M thiourea, 100 mM triethylammonium bicarbonate, a protease/phosphatase inhibitor cocktail, and most importantly 5 % sodium dodecyl sulfate (SDS), was used. The addition of 5% SDS guarantee the extraction of membrane-bound proteins. Indeed, denaturing detergents such as SDS totally disrupt membranes, bind to both membrane (hydrophobic) and non-membrane (water-soluble, hydrophilic) proteins and denature proteins by breaking protein–protein interactions. Non-denaturing detergents such as CHAPS used in the first step have rigid and bulky nonpolar heads that do not penetrate into water-soluble proteins; consequently, they generally do not disrupt membranes or do at low extend³.

Reagents

- Urea (Merck, cat. no. U6504-1KG)

- Thiourea (Merck, cat. no. T8656-500G)
- Triethylammonium bicarbonate (Thermo-Fisher Scientific, cat. no. 90114)
- Protease/phosphatase inhibitor mini tablets (Thermo-Fisher Scientific, cat. no. A32959)
- PlusOne Sodium dodecyl sulfate (SDS; GE Healthcare, cat. no. 17-1313-01)
- 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate (CHAPS; Merck, cat no. C9426-5G)
- Pierce 660 nm Protein Assay Reagent (Thermo-Fisher Scientific, cat. no. 22660)
- 30% Acrylamide/Bis Solution 29:1 (BioRad, cat. no. 1610156)
- Ammonium bicarbonate ³ 99.5% (Merck, cat. no. 09830-500G)
- Acetonitrile Gradient LC-MS (ACN; Scharlab, cat. no. AC03712500)
- Tris(2-carboxyethyl)phosphine hydrochloride solution (TCEP-HCl; Merck, cat. no. 646547)
- PlusOne Dithiothreitol (DTT; GE Healthcare, cat. no. 17-1318-02)
- 2-Chloroacetamide (CAA; Merck, cat. no. 22790-250G-F)
- Iodoacetamide (IAA; Merck, cat. no. I6125-5G)
- Proteomics Grade Trypsin (Merck, cat. no. T6567-5X20UG)
- Pierce MS-Grade Trypsin (Thermo-Fisher Scientific, cat. no. 90057)
- Trifluoroacetic Acid LC-MS Grade (TFA; Thermo-Fisher Scientific, cat. no. 85183)
- Methanol Gradient LC-MS (Scharlab, cat. no. ME032625)
- Ortho-Phosphoric acid 85% HPLC grade (Merck, cat. no. 5.43828)
- Formic acid ³ 96% (FA; Merck, cat. no. 695076-100ML)

Equipment

- TripleTOF 5600 System (SCIEX)
- Q Exactive HF Orbitrap (Thermo Fisher Scientific)
- Mini-PROTEAN® Electrophoresis System (BioRad)

- StageTip C18 Pipette tips 20UL (Thermo Fisher Scientific, cat. no. NC0028381)
- S-Trap™ micro column (Protifi, cat. no. C02-micro-80)
- Proteiner DP Robot (Bruker Daltonics)
- UP50H ultrasonic lab homogenizer (Hielscher Ultrasonics)
- Centrifuge 5804/5804R (Eppendorf)
- 3M™ Empore™ C18 Extraction Disks (Thermo Fisher Scientific, cat. no. 143862)
- Vacuum Concentrator Plus/ (Eppendorf, cat. no. 5305000304)
- Qubit™ Fluorometric Quantitation (Thermo Fisher Scientific)
- Mascot Server v2.7.0.1 (Matrix Science)

Procedure

Microbial cultures.

The binary culture 768-28 (referred to as “the binary culture”) containing “*Candidatus V. moutnovskia*” was used for testing the protocol⁴. Cultures growing robustly on medium with yeast extract and sulphur or sulphate, were used.

I. Proteomic analyses based on non-denaturing detergent.

Cell lysis and In-Gel protein digestion (stacking gel). Cell pellets were dissolved in lysis buffer (8 M urea, 2 M thiourea, 5% CHAPS, 2 mM TCEP-HCl and protease inhibitor). Homogenisation of the cells was achieved by ultrasonication (10 strokes, low amplitude) on ice. After homogenisation, the lysed cells were centrifuged at 20,000×g for 10 min at 4 °C, and the supernatant containing the solubilised proteins was used for LC-MS/MS experiment. Total protein concentration was determined using Pierce 660nm protein assay (Thermo). An aliquot of every sample was diluted with enough loading sample buffer and then applied onto 1.2 cm wide wells of a conventional SDS-PAGE gel (1mm-thick, 4% stacking, and 12% resolving). The run was stopped as soon as the front entered 1 cm into the resolving gel, so that the whole proteome became concentrated in the stacking/resolving gel interface. The unseparated protein bands were visualised by Coomassie staining, excised, cut into cubes (1 mm²), deposited in 96-well plates and processed automatically in a Proteiner DP (Bruker Daltonics, Bremen, Germany). The digestion protocol used was based on⁶⁶ with minor variations: gel plugs were washed firstly, with 50 mM ammonium bicarbonate and secondly with ACN prior to reduction with 10 mM DTT in 25 mM ammonium bicarbonate solution, and alkylation was carried out with 55 mM IAA in 50 mM ammonium bicarbonate

solution. Gel pieces were then rinsed firstly with 50 mM ammonium bicarbonate and secondly with ACN, and then were dried under a stream of nitrogen. Proteomics Grade Trypsin (Sigma Aldrich) at a final concentration of 16 ng/ μ l in 25% ACN/50 mM ammonium bicarbonate solution was added and the digestion took place at 37°C for 4 h. The reaction was stopped by adding 50% ACN/0.5% TFA for peptide extraction. The eluted tryptic peptides were dried by speed-vacuum centrifugation and then desalted onto StageTip C18 Pipette tips (Thermo Scientific) until the mass spectrometric analysis.

Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometric (LC-ESI-MS/MS). A 1 μ g aliquot of each sample was subjected to 1D-nano LC ESI-MSMS analysis using a nano liquid chromatography system (Eksigent Technologies nanoLC Ultra 1D plus, SCIEX, Foster City, CA) coupled to high speed Triple TOF 5600 mass spectrometer (SCIEX, Foster City, CA) with a Nanospray III source. The analytical column used was a silica-based reversed phase Acquity UPLC M-Class Peptide BEH C18 Column, 75 μ m \times 150 mm, 1.7 μ m particle size and 130 Å pore size (Waters). The trap column was a C18 Acclaim PepMapTM 100 (Thermo Scientific), 100 μ m \times 2 cm, 5 μ m particle diameter, 100 Å pore size, switched on-line with the analytical column. The loading pump delivered a solution of 0.1% formic acid in water at 2 μ l/min. The nano-pump provided a flow-rate of 250 nl/min and was operated under gradient elution conditions. Peptides were separated using a 250 minutes gradient ranging from 2% to 90% mobile phase B (mobile phase A: 2% acetonitrile, 0.1% formic acid; mobile phase B: 100% acetonitrile, 0.1% formic acid). The injection volume was 5 μ l.

Data acquisition was performed with a Triple-TOF 5600 System (SCIEX, Concord, ON). Data was acquired using an ion spray voltage floating (ISVF) 2300 V, curtain gas (CUR) 35, interface heater temperature (IHT) 150, ion source gas 1 (GS1) 25, declustering potential (DP) 100 V. All data was acquired using information-dependent acquisition (IDA) mode with Analyst TF 1.7 software (SCIEX, USA). For IDA parameters, 0.25 s MS survey scan in the mass range of 350–1250 Da were followed by 35 MS/MS scans of 100 ms in the mass range of 100–1,800 (total cycle time: 4 s). Switching criteria were set to ions greater than mass to charge ratio (m/z) 350 and smaller than m/z 1,250 with charge state of 2–5 and an abundance threshold of more than 90 counts (cps). Former target ions were excluded for 15s. IDA rolling collision energy (CE) parameters script was used for automatically controlling the CE.

II. Proteomic analyses based on denaturing detergent.

Lysis and protein extraction. Cell pellets were dissolved with multi-chaotropic lysis buffer containing 7 M urea (USB Corporation, Cleveland, OH), 2 M thiourea (Sigma-Aldrich), 5 % SDS (Sigma-Aldrich), 100 mM triethylammonium bicarbonate (TEAB) (Thermo Fisher Scientific) and a protease/phosphatase inhibitor cocktail (Thermo Fisher Scientific). Samples were reduced and alkylated by adding 5 mM TCEP-HCl and

10 mM CAA for 1 hour at 37°C and homogenized by micro tip probe ultrasonication for 1 min on UP50H ultrasonic lab homogenizer (Hielscher Ultrasonics). The homogenate was centrifuged at 16,000 × g for 15 min at 4 °C, and the supernatant containing the solubilized proteins was used for further analysis. Protein concentration was estimated by Pierce 660nm protein assay (Thermo Fisher Scientific).

S-TrapTM Digestion. Protein digestion in the S-Trap filter (Protifi, Huntington, NY, USA) was performed following the manufacturer's procedure with slight modifications. Briefly, 20 µg of protein of each sample was diluted to 40 µL with 5% SDS. Afterwards, 12% phosphoric acid and then seven volumes of binding buffer (90% methanol; 100 mM TEAB) were added to the sample (final phosphoric acid concentration: 1.2%). After mixing, the protein solution was loaded to an S-Trap filter in two consecutive steps, separated by a 2 min centrifugation at 3000 x g. Then the filter was washed 3 times with 150 µL of binding buffer. Finally, 1 µg of Pierce MS-grade trypsin (Thermo-Fisher Scientific) in 20 µL of a 100 mM TEAB solution was added to each sample in a ratio 1:20 (w/w) and spun through the S-Trap prior to digestion. Flow-through was then reloaded to the top of the S-Trap column and allowed to digest in a wet chamber at 37°C overnight. To elute peptides, two step-wise buffers were applied (1) 40 µL of 25 mM TEAB and 2) 40 µL of 80% acetonitrile and 0.2% formic acid in H₂O), separated by a 2 min centrifugation at 3000 x g in each case. Eluted peptides were pooled and vacuum centrifuged to dryness.

LC-ESI-MS/MS. Digested samples were cleaned-up/desalted using Stage-Tips with Empore 3M C18 disks (Sigma-Aldrich). After desalting, peptide concentration was carried out by QubitTM Fluorometric Quantitation (Thermo Fisher Scientific). A 1 µg aliquot of each digested sample was subjected to 1D-nano LC-ESI-MS/MS (Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometric) analysis using a Thermo Easy-nLC 1000 HPLC system (Thermo Fisher Scientific) coupled online to a Q Exactive HF Orbitrap (Thermo Fisher Scientific). Peptides were eluted onto a 50 cm × 75 µm Easy-spray PepMap C18 analytical column at 45°C and were separated at a flow rate of 250 nL/min using a 210 min gradient ranging from 2 % to 95 % mobile phase B (mobile phase A: 0.1% formic acid (FA); mobile phase B: 100 % acetonitrile (ACN), 0.1 % FA). The loading solvent was 2 % ACN) in 0.1 % FA and injection volume was 10 µl.

Data acquisition was performed using a data-dependent top-15 method, in full scan positive mode, scanning 380 to 1800 m/z. Survey scans were acquired at a resolution of 60,000 at m/z 200, with Automatic Gain Control (AGC) target of 3×10^6 and a maximum fill time (IT) of 40 ms. The top 15 most intense ions from each MS1 scan were selected and fragmented via higher energy collisional dissociation (CID). Resolution for HCD spectra was set to 15,000 at m/z 200, with AGC target of 2×10^5 and maximum ion injection time of 120 ms. Isolation of precursors was performed with a window of 2 m/z and the

normalized collision energy (NCE) was 20. Precursor ions with single, unassigned, or six and higher charge states from fragmentation selection were excluded.

MS/MS data analysis and determination of protein abundances.

MS/MS data obtained for individual samples, using either Method I or II, were processed using Analyst® TF 1.7 Software (SCIEX) (for Method I) and Proteome Discoverer v2.4 (Thermo Fisher Scientific) (for Method II). Raw data file conversion tools generated mgf files which were also searched against protein database that included "*Candidatus V. moutnovskia*" and *Thermoproteus uzoniensis* protein sequences from Uniprot/Swissprot Knowledgebase (last update: 2020/03/04, 4.573 sequences) using the Mascot Server v. 2.5.1 (Matrix Science, London, UK) (for Method I) or Mascot Server v2.7.0.1 (Matrix Science, London, UK) (for Method II). Search parameters were set as follows: enzyme, trypsin; allowed missed cleavages, 2; carbamidomethyl (C) as fixed modification and acetyl (Protein N-term), Gln to pyro-Glu (N-term Q), Glu to pyro-Glu (N-term E) and Oxidation (M) as variable modifications. Peptide mass tolerance was set to 25 ppm and 0.05 Da for fragment masses derived from Method I and ± 10 ppm for precursors and 0.02 Da for fragment masses from Method II. The confidence interval for protein identification was set to $\geq 95\%$ ($p < 0.05$) and only peptides with an individual ion score above the 1% False Discovery Rates (FDR) at spectra level were considered correctly identified. The threshold of only one identified peptide per protein identification was used because FDR controlled experiments counter-intuitively suffer from the two-peptide rule⁵. For estimating the FDR, a target-decoy (TD) search strategy is used. The database search for this approach is performed together (concatenated) on the true (target) as well as null (decoy) database. A decoy database is constructed by simply reversing the target database. It is the simplest approach to calculate FDR and assumes that the number of false *Peptide Spectrum Matches* (PSMs) in decoy search (D) will be equal to the number of false PSMs in target search (T) above a given threshold score. Adding up the false hits in decoy and target, the number of false positives is therefore double of the decoy count above threshold. Therefore FDR is calculated as double of the number of false positives divided by the total hits: $FDR = 2 \times D / (T + D)$.

To rank the protein abundance in each sample, the Exponentially Modified Protein Abundance Index (emPAI) was used in the present study as a relative quantitation score of the proteins in a complex mixture based on protein coverage by the peptide matches in a database search result⁶. Although emPAI is not as accurate as quantification using synthesized peptide standards, it is quite useful for obtaining a broad overview of proteome profiles. In this work we used normalized emPAI values (nemPAI), obtained from the emPAI values by dividing each individual value by the sum of all emPAI values in a given experiment⁶. t-Student's tests were used for pairwise comparisons of the relative abundance of proteins of interest using a script in R programming language to automatize the process.

Proteomics data repository.

The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE⁷ partner repository with the dataset identifier PXD012750 and 10.6019/PXD012750.

Troubleshooting

To avoid liquid leakage from the S-Trap column, a customized yellow tip with 9 Empore 3M C18 disks (Sigma-Aldrich) was placed at the bottom tip of the S-Trap column during digestion.

Time Taken

I. Proteomic analyses based on non-denaturing detergent

- Cell lysis and protein extraction, 1 hour
- Determination of protein concentration, 15 min
- SDS-PAGE 'Stacking Gel', 30 min
- In-Gel protein digestion, 7 hours
- Desalting/clean-up onto StageTip C18 Pipette tips, 1 hour
- LC-ESI-MS/MS analysis, 4 hours
- Data analysis and determination of protein abundances, 1 hour

II. Proteomic analyses based on denaturing detergent

- Cell lysis and protein extraction, 1 hour
- Determination of protein concentration, 15 min
- S-TrapTM Digestion, 12 -16 hours
- Desalting/clean-up onto StageTip C18 Pipette tips, 1 hour
- LC-ESI-MS/MS analysis, 3.5 hours
- Data analysis and determination of protein abundances, 1 hour

Anticipated Results

The main purpose of this research was to access a simple protocol by which cytosolic and membrane bound proteins would be possible for the archaeon *Candidatus Vulcanisaeta moutnovskia*, and can be performed with having access to general chemical resources. We have been successful in generating proteomic results from the binary sulphate-reducing archaeal culture 768-28, grown with sulphate and with sulphur as electron acceptors.

References

1. Muers, M. Gene Expression: Transcriptome to proteome and back to genome. *Nat. Rev. Genet.* **12**, 518, (2011).
2. Guo, J. et al. Specialized proteomic responses and an ancient photoprotection mechanism sustain marine green algal growth during phosphate limitation. *Nat. Microbiol.* **3**, 781-790 (2018).
3. Walker, J. M. (2009) *The Protein Protocols Handbook*. Third Edition. New York (NY): Springer-Verlag New.
4. Prokofeva, M. I. et al. Cultivated anaerobic acidophilic/acidotolerant thermophiles from terrestrial and deep-sea hydrothermal habitats. *Extremophiles* **9**, 437-448 (2005).
5. Shevchenko, A. *et al.* Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl. Acad. Sci. USA* **93**, 14440-14445 (1996).
6. Arike, L. & Peil, L. Spectral Counting Label-Free Proteomics. in *Methods in Molecular Biology* (ed. Martins-de-Souza, D.) 213–222 (Media, Springer Science + Business, 2014).
7. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

Acknowledgements

The proteomic analysis was performed in the Proteomics Facility of The Spanish National Center for Biotechnology (CNB-CSIC), which belongs to ProteoRed, PRB2-ISCIII, supported by grant PT13/0001 (to S.C., M.C.M.). M.F. acknowledges the grants PCIN-2017-078 (within the Marine Biotechnology ERA-NET) and BIO2017-85522-R from the Ministerio de Economía y Competitividad, Ministerio de Ciencia, Innovación y Universidades (MCIU), Agencia Estatal de Investigación (AEI), Fondo Europeo de Desarrollo Regional (FEDER) and European Union (EU), and grant ERA NET TRANSCAN-2 AC17/00022 from the Fundación Agencia Española contra el Cáncer and Instituto de Salud Carlos III.